RESULTS

The Delphi survey

- **The Delphi panel:**

     The 25 panelists enrolled in the survey had periods of qualified experience in Dermatology (as calculated from the date of the first specialty graduation) ranging from 2 to 31 years (mean period of experience ± SD =16 ±7.472 years; coefficient of the range = 0.88).

- **Results of the Delphi sets:**

     The following figures **(Figures 3.1-3.5)** and tables **(Tables 3.1-3.5)** embody the results of the Delphi project. In the figures, percentages of accords (the vertical axis) obtained for every rank (the horizontal axis) are represented. In the tables, the nil (rank "0") accords are tabulated against the sum of agreement accords (rank "1" + rank "2" + rank "3") together with the 95% CI for each item results.
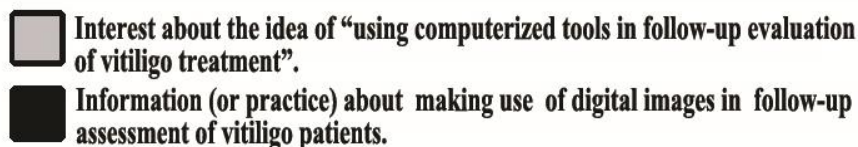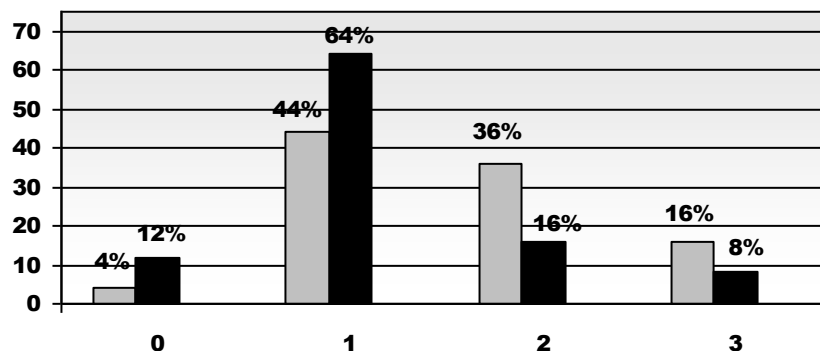
1) <u>**Set 1 items (Figure 3.1 & Table 3.1):**</u>

Interest about the idea of "using computerized tools in follow-up evaluation of vitiligo treatment".

Information (or practice) about making use of digital images in follow-up assessment of vitiligo patients.

<u>**Figure (3.1)**</u> **: Comparison of percentages of panelists' consensuses as regard to the voted ranks in set 1 questionnaire.**

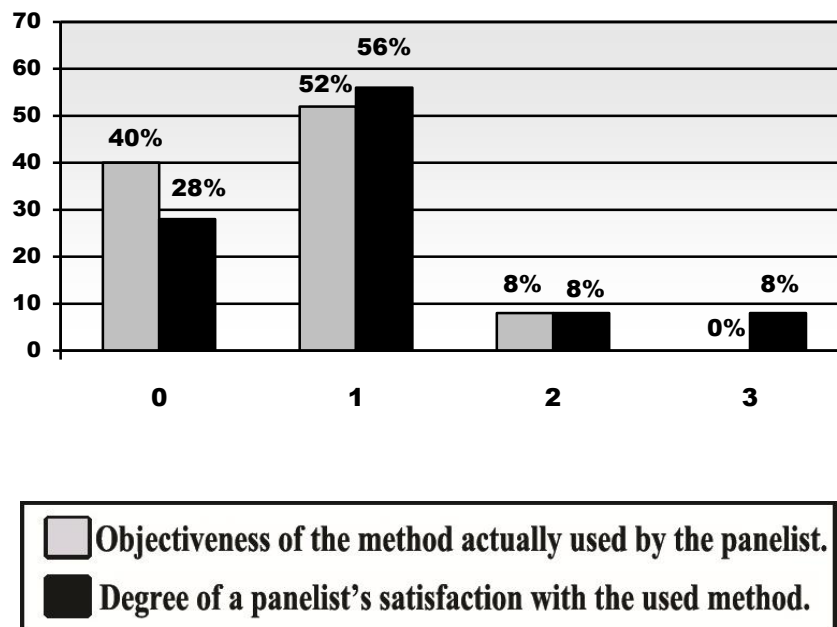| Table (3.1): Comparison of voting accords for rank "0" and those for non-zero ranks in set 1 questionnaire.					
Item	Rank "0"		Non-zero ranks		95% CI
	N	%	N	%	
1	1	4%	24	96%	± 8.28 %
2	3	12%	22	88%	± 13.71 %

2) Set 2 items (Figure 3.2 & Table 3.2):

Figure (3.2): Comparison of the percentages of panelists' consensuses as regard to the voted ranks in set 2 questionnaire.

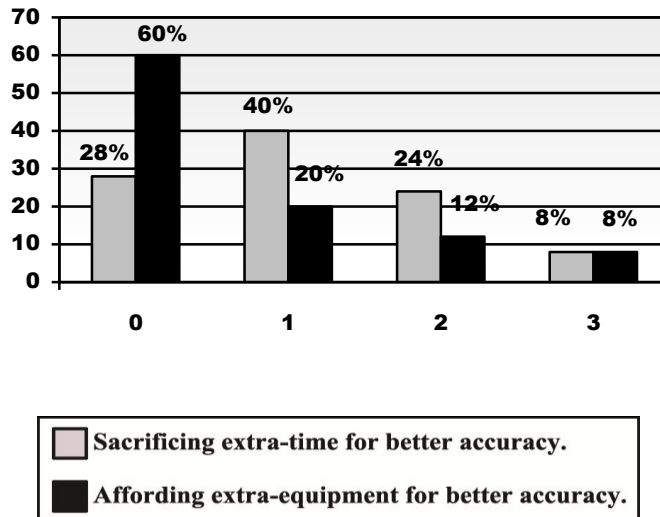| Table (3.2): Comparison of voting accords for rank "0" and those for non-zero ranks in set 2 questionnaire.					
Item	Rank "0"		Non-zero ranks		95% CI
	N	%	N	%	
1	10	40%	15	60%	± 20.66%
2	7	28%	18	72%	± 18.94%

3) Set 3 items (Figure 3.3 & Table 3.3):

Figure (3.3): Comparison of the percentages of panelists' consensuses as regard to the voted ranks in set 3 questionnaire.

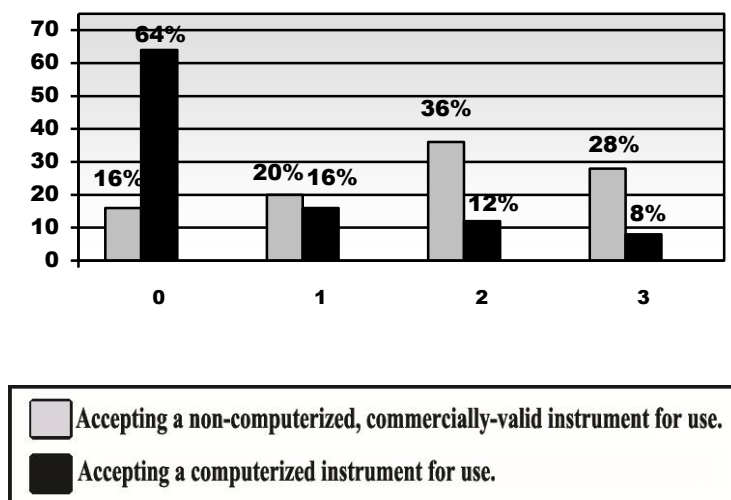| Table (3.3): Comparison of voting accords for rank "0" and those for non-zero ranks in set 3 questionnaire.					
Item	Rank "0"		Non-zero ranks		95% CI
	N	%	N	%	
1	7	28%	18	72%	± 18.94%
2	15	60%	10	40%	± 20.66%

4) Set 4 items (Figure 3.4 & Table 3.4):

Figure (3.4): Comparison of the percentages of panelists' consensuses as regard to the voted ranks in the 2 items targeted in set 4 questionnaire.

Table (3.4): Comparison of voting accords for rank "0" and those for non-zero ranks in set 4 questionnaire.

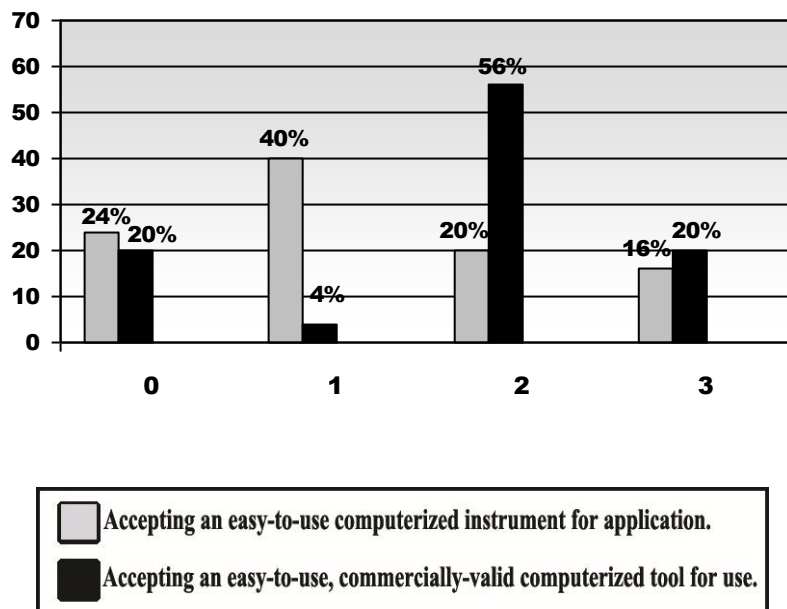| Item	Rank "0"		Non-zero ranks		95% CI
	N	%	N	%	
1	4	16%	21	84%	± 15.47%
2	14	64%	9	36%	± 20.24%

5) Set 5 items (Figure 3.5 & Table 3.5):

Figure (3.5): Comparison of the percentages of panelists' consensuses regarding the voted ranks in the 2 items targeted in set 5 questionnaire.

Table (3.5): Comparison of voting accords for rank "0" and those for non-zero ranks in set 5 questionnaire.

Item	Rank "0"		Non-zero ranks		95% CI
	N	%	N	%	
1	6	24%	19	76%	± 18.01%
2	5	20%	20	80%	± 16.87%

The pilot study

The following items represent quick scanning hints on the results of the pilot study:

1) Pilot study of digital planimetry techniques:

▪ Characteristics of the implicated models:

The circles used in the study were having surface areas ranging from 0.8 to 153.9 cm^2 (mean ±SEM: 46.73 ±16.58; N= 10; coefficient of the range: 0.99).

▪ Clinimetric validation:

1. Accuracy of the obtained SA measurement using DIA$_{SA}$ of scanned traces and of photographed traces

The differences between the measurements obtained by both observers using both methods of digital planimetry and the actual surface areas (gold standard) of circles were found to be statistically non-significant (P >0.05).

2. Inter-observer reproducibility

On comparing the measurements obtained by each observer with those obtained by the other using the same method, significant levels of agreement were revealed (r =1.00, P <0.001) regarding both techniques.

3. Intra-observer reproducibility

Significant agreement levels were found (r =1.00, P <0.001) between measurements obtained by each observer in the 2 trials performed with both digital planimetry techniques.

2) Pilot study of color difference measurement technique:

1. Inter-observer reproducibility

On comparing the results obtained by each 2 observers using one of the 2 shots taken for a single lesion, significant agreement levels were revealed (shot 1: r = 0.998, p <0.001 & shot 2: r = 0.987, P <0.001).

2. Intra-observer reproducibility across changed illumination

Significant agreement levels were shown between color measurements (of the same lesion) obtained in trial '1' (performed on the first shot) and trial '2' (performed on the second shot) by each observer (observer 1: r = 0.908, P <0.001& observer 2: r = 0.912, P <0.001).

Results of analysis of the 3 SA measurement techniques

- **The implicated sample data:**

The measured 100 areas (in both sessions) had a mean SA of 10.16 ±1.85 cm^2 (mean SA ±SEM) **(Table 3.6).**

Table (3.6): Description of the enrolled lesions' actual SAs (in cm^2).

N	Mean	SD	SEM	Range (0.43-119.95 cm^2)	Coefficient of range
100	10.16	18.47	1.85	119.52	0.99

Regarding the enrolled 50 lesions, the periods left between the 2 assessment sessions were ranging from 4 to12 weeks (mean period ±SD = 8.3 ±3.1 weeks) **(Table 3.7).**

Table (3.7): Description of the intervals left between the 2 assessment sessions (in weeks).

N	Mean	SD
50	8.3	3.1

The actual reductions in SAs of lesions occurring after left periods of treatment were ranging from 0.11 to 10.2 cm^2 (mean reduction in SA ±SD = 1.79 ±2.15 cm^2) **(Table 3.8).**

Table (3.8): Description of actual diminutions in SA of the enrolled lesions occurring after the left periods of treatment.

N	Mean	SD	SEM
50	1.79	2.15	0.3

- **Accuracy (Tables 3.9; 3.10 & Figures 3.6; 3.7):**

There were statistically non-significant differences between SA measurements obtained by each observer, using each of the 3 designated methods and the SA measurements constituting the gold standard of accuracy.

Table (3.9): Comparison of the aberrations encountered in measurements obtained using each of the 3 SA measurement techniques as depicted from measurements performed by observer (1).

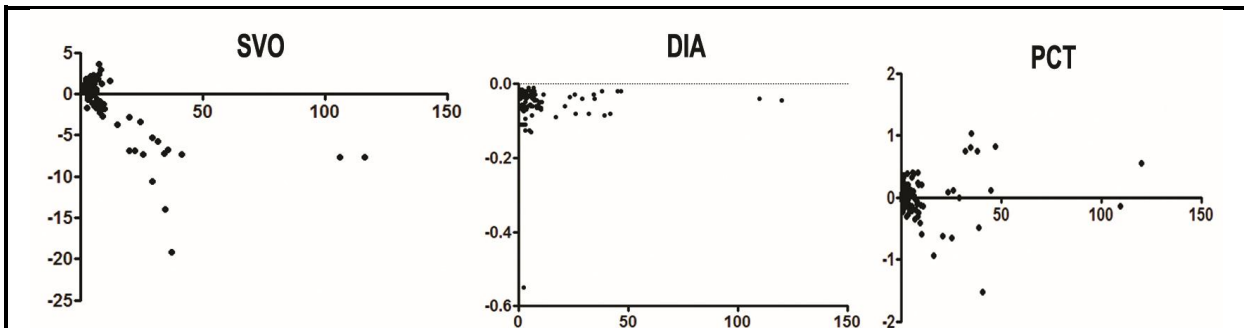| Method	N	Mean	SD	SEM	Sig. of difference		BA test (difference versus average)		
					P	Sig.	Bias	SD of bias	95% Limits of agreement
SVO	100	9.31	16.07	1.61	0.91	NS	-0.86	3.56	From -7.84 to 6.12
DIA$_{SA}$	100	10.11	18.47	1.85	0.41	NS	-0.06	0.05	From -0.16 to 0.06
PCT	100	10.19	18.50	1.85	0.25	NS	0.02	0.34	From -0.64 to 0.68

Figure (3.6): BA plots of biases encountered with SA measurements obtained by observer (1) implying each of the 3 offered methods (using the differences "the vertical axis" versus the average values between the obtained measurements and the gold standard measurements "the horizontal axis").

Table (3.10): Comparison of the aberrations encountered in measurements obtained using each of the 3 SA measurement techniques as depicted from measurements performed by observer (2).

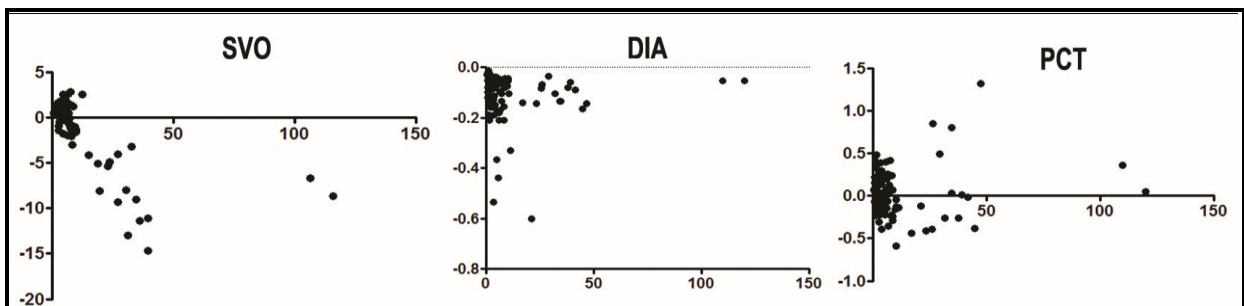| Method	N	Mean	SD	SEM	Sig. of difference		BA test (difference versus average)		
					P	Sig.	Bias	SD of bias	95% Limits of agreement
SVO	100	9.31	16.02	1.60	0.72	NS	-0.86	3.52	From -7.75 to 6.04
DIA$_{SA}$	100	10.16	18.47	1.85	0.38	NS	-0.11	0.1	From -0.30 to 0.08
PCT	100	10.20	18.51	1.85	0.48	NS	0.03	0.28	From -0.51 to 0.58

Figure (3.7): BA plots of biases encountered with SA measurements obtained by observer (2) implying each of the 3 offered methods.

- **Inter-observer reliability (Table 3.11 & Figure 3.8):**

Statistically significant levels of agreement were revealed between measurements obtained by the 2 observers regarding all of the presented methods for SA measurement.

Table (3.11): Comparison of levels of agreement of the measurements obtained by the 2 observers (inter-observer reproducibility) using each of the 3 SA measurement techniques.

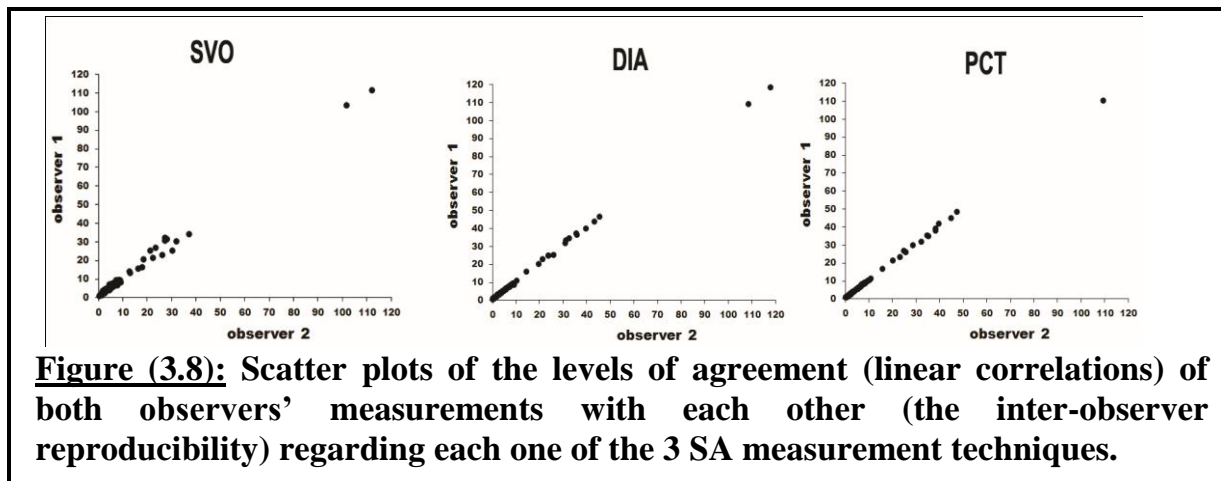| Method	N	Mean		SD		SEM		IC		
		Observer 1	Observer 2	Observer 1	Observer 2	Observer 1	Observer 2	r	P	Sig.
SVO	100	9.31	9.31	16.07	16.02	1.61	1.60	0.9968	<0.0001	S
DIA_{SA}	100	10.11	10.16	18.47	18.47	1.85	1.85	1	<0.0001	S
PCT	100	10.19	10.20	18.50	18.51	1.85	1.85	0.9998	<0.0001	S

Figure (3.8): Scatter plots of the levels of agreement (linear correlations) of both observers' measurements with each other (the inter-observer reproducibility) regarding each one of the 3 SA measurement techniques.

- **Intra-observer reliability (Tables 3.12; 3.13 & Figures 3.9; 3.10):**

Statistically significant levels of agreement were revealed between measurements obtained by the each observer in the 2 performed trials regarding the 3 tested methods for SA measurement.

Table (3.12): Comparison of levels of agreement of the measurements obtained in the 2 trials performed by observer (1) using each SA measurement technique.

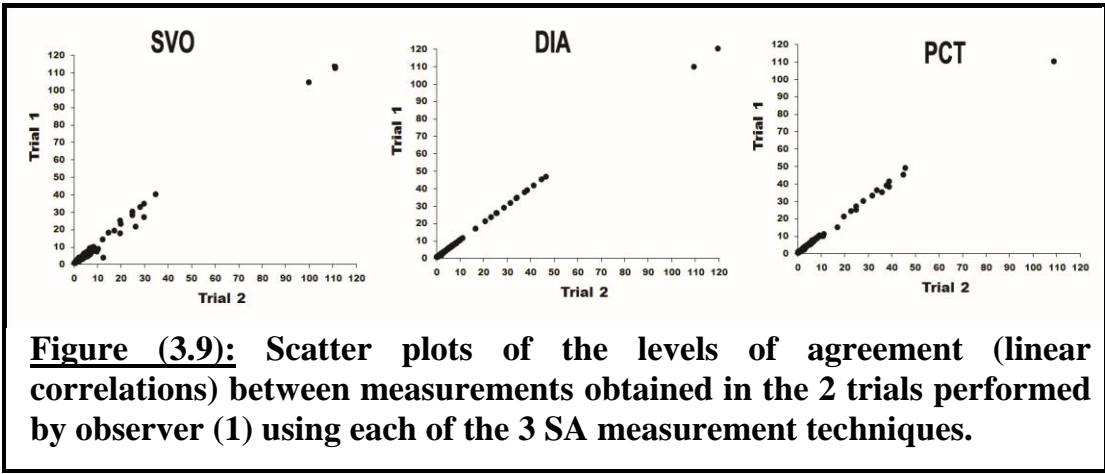| Method	N	Mean		SD		SEM		IC		
		Trial 1	Trial 2	Trial 1	Trial 2	Trial 1	Trial 2	r	P	Sig.
SVO	100	9.17	9.35	15.70	16.48	1.57	1.65	0.9925	<0.0001	S
DIA_{SA}	100	10.12	10.11	18.47	18.47	1.85	1.85	1	<0.0001	S
PCT	100	10.14	10.23	18.48	18.53	1.85	1.85	0.9992	<0.0001	S

Figure (3.9): Scatter plots of the levels of agreement (linear correlations) between measurements obtained in the 2 trials performed by observer (1) using each of the 3 SA measurement techniques.

Table (3.13): Comparison of levels of agreement of the measurements obtained in the 2 trials performed by observer (2) using each SA measurement technique.

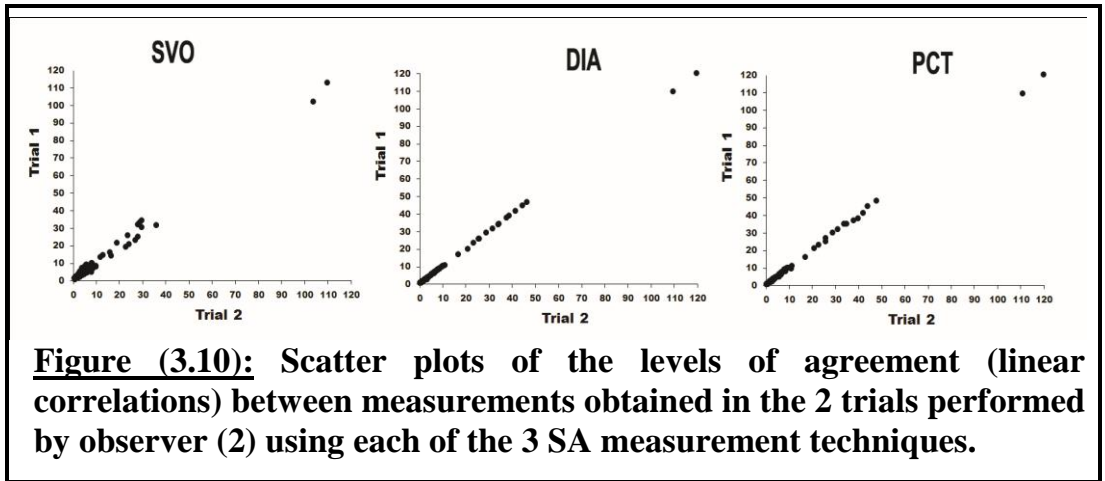| Method	N	Mean		SD		SEM		ICC		
		Trial 1	Trial 2	Trial 1	Trial 2	Trial 1	Trial 2	r	P	Sig.
SVO	100	9.22	9.38	16.04	16.04	1.60	1.60	0.9944	<0.0001	S
DIA$_{SA}$	100	10.07	10.04	18.47	18.46	1.85	1.85	1	<0.0001	S
PCT	100	10.23	10.17	18.56	18.45	1.85	1.86	0.9996	<0.0001	S

Figure (3.10): Scatter plots of the levels of agreement (linear correlations) between measurements obtained in the 2 trials performed by observer (2) using each of the 3 SA measurement techniques.

- **Sensitivity:**

The biases encountered in SA alterations occurring in the second assessment session as obtained by observer '1' (**Table 3.14 & Figure 3.11**) were (mean bias ±SD of bias): -0.37 ±1.14 with using SVO; 0.034 ±0.074 with using DIA$_{SA}$; and 0.009 ±0.74 with using PCT.

Table (3.14): Comparison between measured changes in SA of lesions in the second assessment session using the 3 SA measurement techniques as depicted from measurements performed by observer (1).

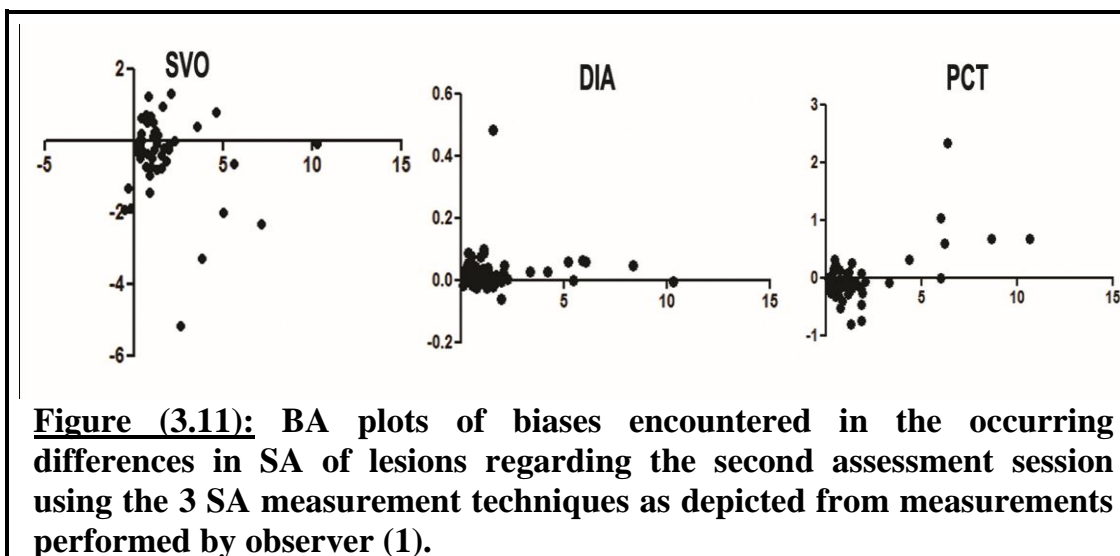| Method	N	Mean	SD	BA test (difference versus average)		
				Bias	SD of bias	95% Limits of Agreement
SVO	50	1.42	1.95	-0.37	1.14	From -2.6 to 1.86
DIA$_{SA}$	50	1.82	2.15	0.034	0.074	From -0.11 to 0.18
PCT	50	1.80	2.44	0.009	0.47	From -0.91 to 0.93

Figure (3.11): BA plots of biases encountered in the occurring differences in SA of lesions regarding the second assessment session using the 3 SA measurement techniques as depicted from measurements performed by observer (1).

Considering measurements performed by observer '2' (**Table 3.15 & Figure 3.12**), the met biases were: -0.40 ±1.34 with using SVO; 0.03 ±0.12 with using DIA$_{SA}$; and 0.02 ±0.40 with using PCT.

Table (3.15): Comparison between measured changes in SA of lesions in the second assessment session using the 3 SA measurement techniques as depicted from measurements performed by observer (2).

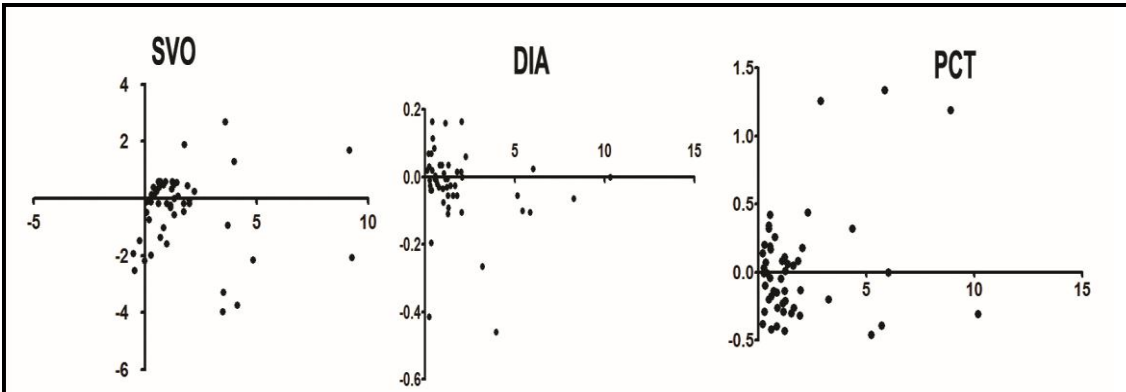| Method	N	Mean	SD	BA test (difference versus average)		
				Bias	SD of bias	95% Limits of Agreement
SVO	50	1.39	2.09	-0.40	1.34	From -3.03 to 2.24
DIA$_{SA}$	50	1.76	2.13	-0.03	0.12	From -0.26 to 0.20
PCT	50	1.81	2.26	0.02	0.40	From -0.76 to 0.79

Figure (3.12): BA plots of biases encountered in the occurring differences in SA of lesions regarding the second assessment session using the 3 SA measurement techniques as depicted from measurements performed by observer (2).

Results of analysis of the obtained % of treatment results calculated from color difference measurement using DIA_{CM}

- **Inter-observer reliability** (Table 3.16 & Figure 3.13):

A statistically significant level of agreement was revealed (r = 0.995, P <0.001) between percentages of treatment results in the used lesion models obtained by the 2 observers regarding the offered color difference measurement technique.

Table (3.16): Comparison of % of treatment results obtained by the 2 observers using DIA_{CM}.

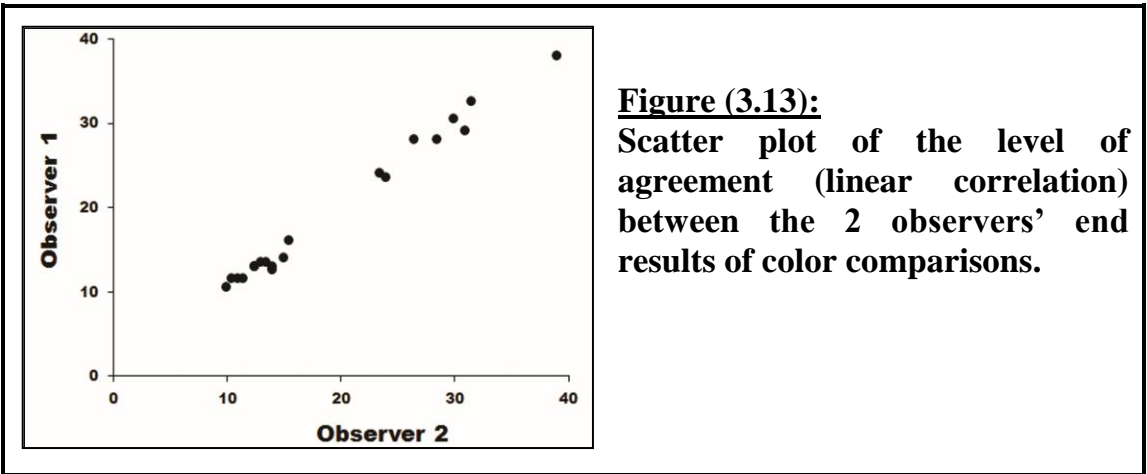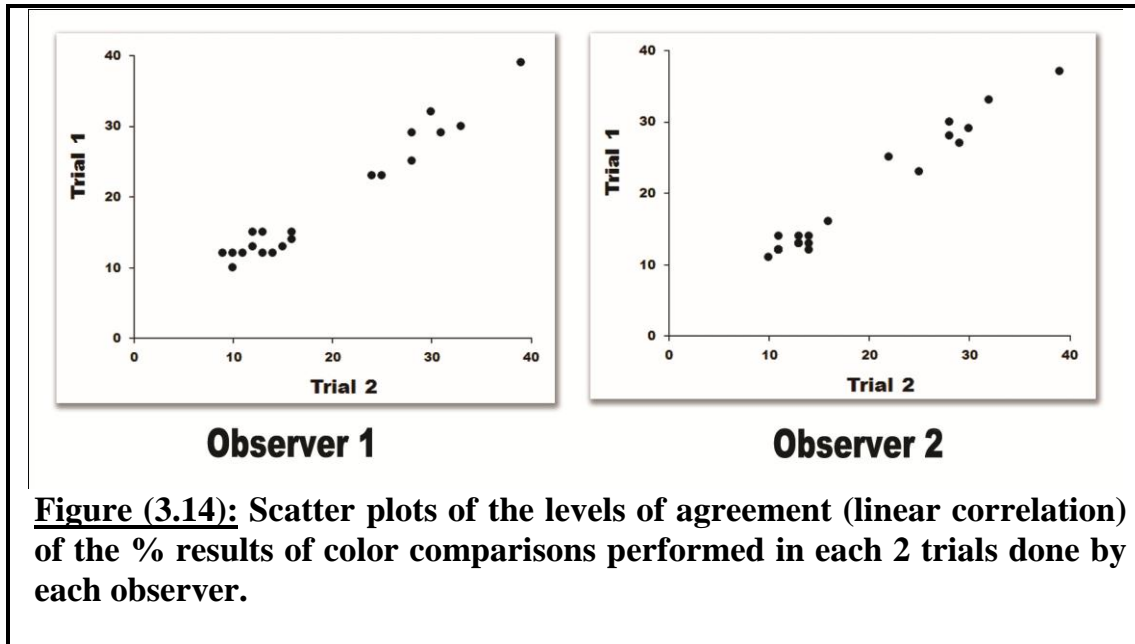| N	Mean		SD		SEM		ICC		
	Observer 1	Observer 2	Observer 1	Observer 2	Observer 1	Observer 2	r	P	Sig.
20	19.35	19.35	8.92	8.78	2	1.96	0.995	<0.001	S

Figure (3.13):
Scatter plot of the level of agreement (linear correlation) between the 2 observers' end results of color comparisons.

- **Intra-observer reliability** (Table 3.17 & Figure 3.14):

Statistically significant levels of agreement were revealed between percentages of treatment results obtained in the 2 trials done by each observer (r = 0.977 for observer 1; 0.985 for observer 2, P <0.001 for both observers) on implying the offered color difference measurement technique on the selected lesion models.

Table (3.17): Comparison of % of treatment results obtained by each observer in the performed 2 trials using DIA$_{CM}$.

observer	N	Mean		SD		SEM		ICC		
		Trial 1	Trial 2	Trial 1	Trial 2	Trial 1	Trial 2	r	P	Sig.
1	20	19.45	19.25	9.28	8.67	2.07	1.94	0.977	<0.001	S
2	20	19.20	19.40	8.96	8.54	2.00	1.91	0.985	<0.001	S

Figure (3.14): Scatter plots of the levels of agreement (linear correlation) of the % results of color comparisons performed in each 2 trials done by each observer.

- **Simulated accuracy** (Table 3.18 & Figure 3.15):

A statistically significant level of agreement was revealed between percentages of treatment results obtained on assessing lesions selected from bilaterally symmetrical body portions (P<0.001; r = 0.962).

Table (3.18): Comparison of % of treatment results obtained in each 2 bilateral lesions using DIA$_{CM}$.

N	Mean		SD		SEM		ICC		
	Lesion 1	Lesion 2	Lesion 1	Lesion 2	Lesion 1	Lesion 2	r	P	Sig.
10	19.85	18.88	9.98	8.09	3.16	2.56	0.962	<0.001	S
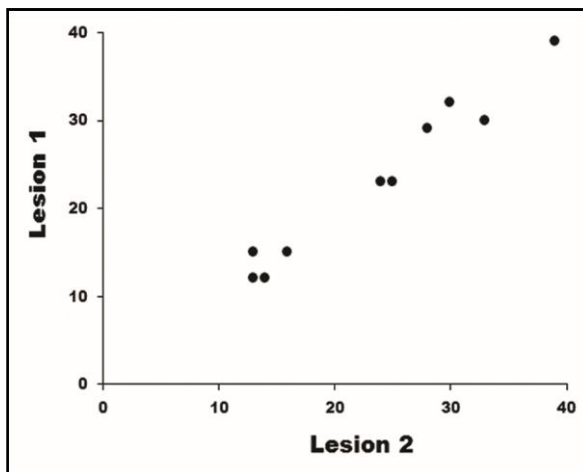
Figure (3.15):
Scatter plot of the level of agreement (linear correlation) between the responses shown by each 2 bilaterally symmetrical lesions.